Vaccine Safety Datalink Project: A New Tool for Improving Vaccine Safety Monitoring in the United States

Robert T. Chen, MD, MA*; John W. Glasser, PhD, MPH*; Philip H. Rhodes, PhD*; Robert L. Davis, MD, MPH‡; William E. Barlow, PhD‡; Robert S. Thompson, MD‡; John P. Mullooly, PhD§; Steve B. Black, MD‖; Henry R. Shinefield, MD‖; Connie M. Vadheim, PhD¶; S. Michael Marcy, MD#; Joel I. Ward, MD¶; Robert P. Wise, MD, MPH**; Steven G. Wassilak, MD*; Stephen C. Hadler, MD*; and the Vaccine Safety Datalink Team‡‡

ABSTRACT. *Objective.* **To fill the large "gaps and limitations" in current scientific knowledge of rare vaccine adverse events identified in recent reviews of the Institute of Medicine.**

Methods. **Computerized information on immunization, medical outcomes, and potential confounders on more than 500 000 children 0 to 6 years of age is linked annually at several health maintenance organizations to create a large cohort for multiple epidemiologic studies of vaccine safety.**

Results. **Analysis of 3 years of follow-up data shows that 549 488 doses of diphtheria-tetanus-pertussis (DTP) and 310 618 doses of measles-mumps-rubella (MMR) vaccines have been administered to children in the study cohort. Analyses for associations between vaccines and 34 medical outcomes are underway. Screening of automated data shows that seizures are associated with receipt of DTP on the same day (relative risk [RR], 2.1; 95% confidence interval [CI], 1.1 to 4.0) and 8 to 14 days after receipt of MMR (RR, 3.0; 95% CI, 2.1 to 4.2). The diversity of vaccination exposures in this large cohort permits us to show that an apparent association of seizures 8 to 14 days after *Haemophilus influenzae* type b vaccine (RR, 1.6; 95% CI, 1.2 to 2.1) was attributable to confounding by simultaneous MMR vaccination; the association disappears with appropriate adjustment (RR, 1.0; 95% CI, 0.7 to 1.4).**

Conclusion. **Preliminary design, data collection, and analytic capability of the Vaccine Safety Datalink project has been validated by replication of previous known associations between seizures and DTP and MMR vaccines. The diversity in vaccine administration schedules permits potential disentangling of effects of simultaneous and combined vaccinations. The project provides a model of public health-managed care collaborations in addition to an excellent infrastructure for safety and other studies of vaccines.** *Pediatrics* **1997;99:765–773;** ***vaccines, immunization, adverse reactions, databases, record linkage, vaccine safety.***

From the *National Immunization Program, Centers for Disease Control and Prevention, Atlanta, Georgia; ‡Center for Health Studies, Group Health Cooperative of Puget Sound, Seattle, Washington; §Center for Health Research, Northwest Kaiser Permanente, Portland, Oregon; ‖Pediatric Vaccine Study Center, Northern California Kaiser Permanente, Oakland; ¶Center for Vaccine Research, Harbor-UCLA Medical Center, Torrance, California; #Southern California Kaiser Permanente, Pasadena; and **Center for Biologics Evaluation and Research, Food and Drug Administration, Rockville, Maryland.

‡‡Other Vaccine Safety Datalink team members by site: Emmett Swint, MA*; Janet R. Hardy, MSc, MPH*; Thomas Payne, MD‡; Virginia Immanuel, MPH‡; Patti Benson‡; Julie Drake‡; Lois Drew§; Barbara Mendius, MS§; Paula Ray, MPH‖; Ned Lewis, MPH‖; Bruce H. Fireman, MA‖; Jennie Jing, MA¶; Michael Wulfsohn, MD, PhD¶; Marlene M. Lugg, DrPH#; Patricia Osborne, MS#; Suresh Rastogi, PhD**; Peter Patriarca, MD**; and Vito Caserta, MD, MPH (Division of Vaccine Injury Compensation, Health Resources and Services Administration, Rockville, Maryland).

Received for publication Apr 1, 1996; accepted Jun 14, 1996.

Reprint requests to (R.T.C.) National Immunization Program (MS-E61), Centers for Disease Control and Prevention, Atlanta, GA 30333.

ABBREVIATIONS. VAERS, Vaccine Adverse Event Reporting System; IOM, Institute of Medicine; CDC, Centers for Disease Control and Prevention; LLDB, large linked database; VSD, Vaccine Safety Datalink; HMO, health maintenance organization; GHC, Group Health Cooperative (of Puget Sound); NWK, Northwest Kaiser; NCK, Northern California Kaiser; SCK, Southern California Kaiser; DTP, diphtheria-tetanus-pertussis; RR, relative risk; ICD, International Classification of Diseases; Hib, *Haemophilus influenzae* type b; MMR, measles-mumps-rubella; CI, confidence interval; Td, tetanus-diphtheria.

Immunizations are among the most cost-effective and widely used public health interventions.[1] The incidence rates of vaccine-preventable diseases in the United States and most countries worldwide[2] have decreased dramatically during recent decades (Table 1). No vaccine is perfectly safe, however. Increased vaccine use necessarily results in an increased number of true vaccine reactions as well as adverse medical events coincidentally associated with vaccinations.[3] The number of both types of reports to the Vaccine Adverse Event Reporting System (VAERS) in the United States, approximately 10 000 per year, now exceeds the reported incidence of most vaccine-preventable childhood diseases combined (Table 1).

Because few vaccine-preventable diseases are currently eradicable, most immunizations must be continued indefinitely. One important way to minimize vaccine injuries is to improve our understanding of vaccine safety and thereby foster the development and use of safer vaccines.[4] Close monitoring of vaccine safety should also help prevent the loss of public confidence in immunization programs and the subsequent resurgence of vaccine-preventable diseases, as experienced with pertussis in several countries[5–7] and more recently with diphtheria.[8]

Despite the importance of vaccine safety, the Institute of Medicine (IOM) recently found that serious "gaps and limitations" exist in both the knowledge

TABLE 1. Comparison of Maximum and Current Reported Morbidity From Vaccine-preventable Diseases and Vaccine Adverse Events, United States

Disease	Pre-Vaccine Era		1996*	Change, %
	Maximum Cases	Year		
Diphtheria	206 939	(1921)	1	−99.99
Measles	894 134	(1941)	488	−99.95
Mumps	152 209	(1968)	658	−99.57
Pertussis	265 269	(1934)	6467	−97.56
Polio (wild)	21 269	(1952)	0	−100.00
Rubella	57 686†	(1969)	210	−99.64
Congenital rubella syndrome	20 000†	(1964–1965)	2	−99.99
Tetanus	1 560†	(1948)	27	−98.27
Haemophilus influenzae type b invasive disease	20 000	(1984)	276	−98.62
Vaccine adverse events	0		11 690	††

* Provisional, subject to change because of late reporting.
† Estimated because no national reporting existed in the prevaccine era.
†† This indicates the major increase in vaccine adverse events.

and infrastructure needed to study vaccine adverse events.[9,10] Among the 76 types of vaccine adverse events reviewed by the IOM, the scientific evidence was inadequate to assess definitive vaccine causality for 50 (66%). The IOM also noted that "if research . . . [is] not improved, future reviews of vaccine safety will be similarly handicapped." These gaps in knowledge are attributable to several factors. Prelicensure controlled trials provide only limited safety data because of their relatively small sample size, short duration, and population homogeneity. Postlicensure studies are therefore needed to provide a fuller understanding of the safety of vaccines in general use.[4]

Historically, postlicensure studies of safety have relied on passive surveillance systems such as the VAERS.[3] Because of methodological weaknesses, such as the potential for biased reporting and underreporting and lack of denominators or comparison groups, data from such case reports are usually not helpful in assessing risk or vaccine causality.[4] Ad hoc retrospective epidemiologic studies of vaccine safety (eg, swine influenza vaccine and Guillain-Barré syndrome[11] and pertussis vaccine and encephalopathy[5]), although potentially informative about vaccine causality, are costly, time consuming, and usually limited to assessment of a single event.

Recognizing the need to improve the capability to study vaccine safety, the Centers for Disease Control and Prevention (CDC) participated during the late 1980s in two pilot studies using large linked databases (LLDBs) of computerized vaccination and medical records.[12–14] These studies helped validate the LLDB approach for vaccine safety studies. The need for a larger LLDB population for continuous assessment of vaccine safety prompted the CDC to initiate planning for the Vaccine Safety Datalink (VSD) project in 1989.[15] This article provides an overview of the design of the VSD project and preliminary results and reviews the prospects and the limitations of the VSD to address the vaccine safety issues identified by the IOM, as well as those of future vaccines.[16]

VSD PROJECT

Background

The need to improve postlicensure monitoring of drug safety became widely recognized after the thalidomide disaster.[17] Faced

TABLE 2. Example of Method for Risk-interval Analysis of an Association Between a Universally-recommended 3-Dose Vaccine (With Few Unvaccinated Persons for Comparison) and Adverse Event: Vaccine Safety Datalink Project

1. Define "risk interval" for adverse event after vaccination (eg, 30 days after each dose).
2. Partition observation time for each child in the study into periods within and outside of risk intervals, and sum respectively (eg, for a child observed for 365 days during which 3 doses of vaccine were received; total risk interval time = 3 × 30 person-days = 90 person-days; total non–risk interval time = 365 − 90 = 275 person-days).

o——————x-----——————x-----——————x-----——————→
Birth Dose 1 Dose 2 Dose 3

3. Add up: (a) total risk interval and non–risk interval observation times for each child in the study (=person-time Observed; for mathematical convenience, example below uses 100 and 1000 person-months of observation); and (b) adverse events occurring in each period to complete 2 × 2 table (for illustration, example below uses 3 and 10 cases):

	Adverse Event, Yes	Person-Time Observed, mo	Incidence Rate
Vaccinated in risk interval, yes----	3	100	0.03
Vaccinated in risk interval, no——	10	1000	0.01
Total	13	1100	

Incidence rate adverse event (vaccinated) = 3/100 = 0.03
Incidence rate adverse event (unvaccinated) = 10/1000 = 0.01
Relative risk vaccinated:unvaccinated = 0.03/0.01 = 3.0
Probability finding due to chance: <5/100
Conclusion: There is a 3-fold increase in risk for development of the adverse event within the interval after vaccination.

with the methodological limitations in passive surveillance for drug adverse events, pharmacoepidemiologists began during the 1980s to turn to LLDBs linking computerized pharmacy prescription and medical outcome records.[18] These databases derive from defined populations, such as members of health maintenance organizations (HMOs), single-provider health care systems, and Medicaid programs. Because the databases are usually generated in the routine administration of such programs and do not require completion of an adverse event reporting form, the problems of underreporting or recall bias are reduced. Because these programs have enrollees numbering from thousands to millions, large populations can be examined for relatively infrequent adverse events. Denominator data on doses administered and the ready availability of appropriate comparison groups are also very useful. Therefore, LLDBs can potentially provide an economical and rapid means of conducting postlicensure studies of the safety of drugs and vaccines (Table 2).[19,20]

Study Sites and Population

In 1991, the CDC began a partnership with four HMOs to evaluate vaccine safety in children in a large-scale prospective study. The HMOs include Group Health Cooperative (GHC) of Puget Sound in Washington, Northwest Kaiser Permanente (NWK) in Oregon, Northern California Kaiser (NCK), and Southern California Kaiser (SCK) Permanente health programs. These sites were chosen for their research experience and the existence of or willingness to create HMO-wide computerized vaccination databases for this project. Other Public Health Service agencies, including the Center for Biologics Evaluation and Research of the Food and Drug Administration and the Division of Vaccine Injury Compensation of the Health Resources and Services Administration, have been important contributors to this ongoing project.

The initial study focused on children 0 to 6 years of age but is being expanded to include adolescents and adults. To be eligible for the study, each child must be a member of the HMO, be within predefined age limits, live within the catchment area of a given site's participating clinics, and receive vaccine in a study clinic. The power to examine any particular vaccine and adverse event association within the VSD depends not only on the frequency of the vaccination but also on the sensitivity and specificity of the case definition, the background rate for the event, and the magnitude of the risk. Previous LLDB studies reviewed adverse event experience after approximately 500 000 doses of diphtheria-tetanus-pertussis (DTP) vaccine in 10 years. Our study accumulates similar exposure experience after 2 years. With time, we expect that the VSD study should have adequate power to detect potential vaccine reactions with attributable risk of about 1 outcome per 100 000 vaccine doses (assuming a three-dose vaccine, relative risk [RR] of 1.6, and background rate of 1.2×10^{-5} to 1.6×10^{-3}). This magnitude of attributable risk is similar to that suggested by acute encephalopathy after whole-cell pertussis vaccination[5,9] or Guillain-Barré syndrome after swine influenza vaccination.[11]

Data Collection

Health service use information for each patient is computerized and continuously compiled by each HMO indexed by an unique identifier. These data were initially used primarily for internal HMO administrative and clinical patient treatment purposes but have been adapted to this study. Data collection in a standardized VSD format began on March 1, 1991, for three HMOs (GHC, NWK, and NCK) and October 1, 1992, for SCK. The data are organized into files containing demographic information, covariate information, vaccination records, and various types of medical outcome data (Table 3). The automated outcome data are collected from various sources at each site, such as records of hospitalizations (all sites), emergency department visits (all sites), and outpatient clinic visits (GHC, NWK, and NCK). Each site encodes their patients' clinical data with unique study identifiers before shipping the data to the CDC annually for merging and analysis, thereby preserving patient confidentiality. Institutional review boards at each HMO have approved this project, in which only analyses with aggregate data are presented.

Vaccination Records

All vaccinations given within the HMO study population, either routinely or for special indications, are recorded and entered into a computer database. For all patients, automated data include: (1) the vaccine type, (2) the date of vaccination, and (3) concurrent vaccinations. For most patients, additional automated data available include whether vaccinations were obtained in or outside the HMO, as well as information required by the National Childhood Vaccine Injury Act of 1986[21]: (1) the manufacturer, (2) lot number, and (3) site of vaccination.

TABLE 3. Data Files Created for Vaccine Safety Datalink Project

File Name	Description/Content*
Essential information	
Constant	Unique identifier, birth date, gender
Enrollment	Start and stop dates for enrollment in the study, reasons for leaving the study
Vaccine	Immunization records, vaccine type, and date of administration
Outcome	Hospital and emergency department visits (all sites) + outpatient clinic visits (2 sites)
Ancillary information (outcomes)	
Procedure	Selected procedures (eg, computed tomographic scans and magnetic resonance images)
Laboratory	Selected results of pathogen-specific cultures and other diagnostic tests
Pharmacy	Drug use by classification (eg, anticonvulsants)
Covariate data	
Geocode	For estimating socioeconomic status based on census block codes
Birth	Birth certificates for covariate determination via the geocode file
Deaths	Review of state death certificates and chart review
Past medical diagnoses	International Classification of Diseases, 9th revision, diagnostic codes for prior hospitalizations

* Each study site obtains the necessary information for files from unique, site-specific administrative databases for health care delivery.

Adverse Medical Events (Outcomes) of Interest

International Classification of Diseases (ICD)-9 codes for all hospitalizations and emergency department visits are compiled for the VSD study cohort. Automated diagnostic codes for routine outpatient clinic visits are currently available on approximately half of the cohort. This proportion will increase substantially as the HMOs switch to automated outpatient record systems. For initial study, the project identified 34 principal medical outcomes of possible association with vaccinations (Table 4) based on a thorough review of the literature and the IOM reports. In addition, the safety of certain vaccination practices are to be evaluated. These include simultaneous versus combined vaccination of various antigens and the relative adequacy of observation of contraindications to vaccinations. The list of research questions is amended yearly based on evolving issues in the medical literature and new vaccination policy needs.

The 34 outcomes are identified in the medical outcomes files by one or more ICD-9 diagnostic codes. Additional outcome information is obtained from selected ancillary sources (eg, procedures and laboratory tests). These sources may be used both for case finding and to supplement primary diagnostic codes. For example, positive blood cultures may enhance the accuracy of a discharge diagnosis of bacteremia. The case definitions for each of these 34 outcomes, consisting of one or more different ICD-9 and other diagnostic codes depending on source, have been determined iteratively and continue to evolve with the review of each year's analyses.

As a potential adverse event of immunization, death occurring soon (eg, 60 days) after vaccination is important to assess. However, deaths in children often occur in the home or outside of the routine health care system (eg, at an accident site or as a result of sudden infant death syndrome). Therefore, we maintain surveillance of state death reports in addition to monitoring hospitalizations and emergency department visits within the HMO. State death records are linked to LLDB cohort patients using a probabilistic matching algorithm.[22] Once identified, all deaths in study

TABLE 4. Outcomes of Primary Interest for Initial Evaluation, Vaccine Safety Datalink Project

Category	Outcome of Interest
Neurologic	Aseptic meningitis
	Idiopathic increased intracranial pressure
	Encephalitis/encephalopathy
	Ataxia
	Seizures and persistent seizure disorders
	Reye's syndrome
	Transverse myelitis
	Guillain-Barré syndrome
	Cranial nerve disorders
	Peripheral nerve disorders
	Hearing loss
	Polio and acute paralytic syndromes
Allergic	Allergic reactions, including anaphylaxis
	Asthma/bronchitis
Hematologic	Hemolytic anemia
	Thrombocytopenia
Infectious/ inflammatory	Diarrhea
	Invasive bacterial disease
	Autoimmune/immune complex diseases
	Vaccine-preventable diseases
	Nonbacterial pneumonia
	Myocarditis
	Pancreatitis
	Parotitis
	Arthropathy/arthritis
Metabolic	Hypoglycemia
	Diabetes
Other	Site abscesses
	Persistent crying
	Collapse-hypotonic hyporesponsive episodes
	Breath holding
	Sudden infant death syndrome/unexpected death
	Apnea
	Vaccine adverse events
Practices	Simultaneous/combined vaccinations
	Observation of contraindications

cohort patients are further validated by review of medical and autopsy records.

Potential Covariates

Factors that affect both vaccination status and incidence of medical events may confound observational studies of vaccine safety.[23] Therefore, we obtain from state birth certificate tapes relevant information, such as parental education and occupation, birth weight, and Apgar score. Similarly, additional information on socioeconomic status is obtained on the VSD study cohort by linking the zip codes and street addresses of the patients with their respective census tract blocks via "geocode."[24]

Data Quality Control Procedures

To study potential rare associations between vaccines and adverse events, large and accurate databases are needed.[25] Routine procedures for assuring the quality of HMO databases vary by type of database and HMO. Inpatient databases are important for hospital and financial management; therefore, data quality is high because of staff training, standardized coding protocols, reliability monitoring, and routine audits. Both routine data quality audits required by hospital accrediting agencies and special audits conducted by other research projects have found inpatient databases to be complete and accurate.[26] Drug, laboratory, radiologic, and referral databases contain user-based data that involve clinical and administrative data essential for delivery of medical services. Such information is constantly monitored for accuracy and is used in other research projects.[27]

In addition to routine quality checks for each of the databases, a random 2% sample of the study populations (1% at the larger sites: NCK and SCK) is selected periodically to review the quality of automated vaccination and diagnostic data. For these samples, vaccination and diagnostic data are abstracted from the medical records and compared with the automated data. Preliminary data from the first year of the VSD found nearly perfect agreement between automated and abstracted hospital discharge diagnoses.[28] The percentage agreement between the dates of the abstracted and automated emergency departments diagnoses ranged from 60% to 87%. The percentage agreement between abstracted and automated vaccination dates ranged from 70% to 98% across vaccines (with the exception of DT) and HMOs. The primary source of disagreement was the incomplete entry of all vaccinations into the database. Continuous feedback to HMOs is provided to identify potential means of improving the quality of the automated data.

Analytic Strategy

Inference that a vaccine causes an adverse event may be drawn if higher rates are consistently observed among vaccinees compared with nonvaccinees, especially if results are consistent for the different HMOs and for several study years. Because of high vaccine coverage within the HMOs, for most outcomes of interest, there are too few unvaccinated persons, and those not vaccinated may not constitute representative comparison group for epidemiologic analyses. Alternatively, an analysis of "risk interval" is used (Table 2). For each outcome of interest, time intervals after vaccination are selected during which an adverse event would be expected if an association exists. These intervals are defined a priori based on biological and clinical considerations. For the study population, incidence rates of adverse events within and outside the specific risk interval for the study population are then compared using appropriate statistical methods, after controlling for potential confounders.

Several statistical analytical designs are used in the VSD study to compare incidence rates, depending on the accuracy and completeness of the automated data. If the quality of the automated data for an outcome of interest is high, multivariate cohort methods (eg, Poisson or proportional hazards regression) are used. Otherwise, the automated records will require validation by medical record review for all persons suspected of having the outcome. Similar validation of the medical records of the random 1% to 2% of the study population selected for quality control assessment permits a more efficient means of providing the comparison group for a case-cohort analysis.[29] Validation studies using nested case-crossover[30] or case-control methods may also be done in parallel or sequentially, especially if a previously unknown association is detected. The medical records of cases, or cases and controls (also selected from the 1% to 2% samples if possible), will be thoroughly reviewed, and children or their parents may be interviewed to identify other potential confounders.

RESULTS

Descriptive Epidemiology

Because each HMO has its own set of medical and operating procedures, much of the first 2 years of the VSD study was devoted to developing protocols to standardize prospective data collection across sites. NCK and SCK incrementally expanded their automated vaccination registries throughout the HMOs. Consequently, the annual study population under surveillance grew from approximately 181 000 in 1991 to approximately 502 000 children younger than 7 years (approximately 2% of the US population in these age groups) by late 1996, encompassing 1 862 000 child-years of observation. Complete vaccination and other medical records are available now on a cohort of 242 000 children born into the study, which will grow by about 63 700 children annually. The 10 most common vaccines and vaccine combinations administered to the cumulative study cohort by late 1994 are shown in Table 5.

Figure 1 plots the vaccine coverage for specific vaccines by age for the VSD cohort. The delay between the recommended age for a vaccine dose and

TABLE 5. Ten Most Common Vaccine and Vaccine Combinations Administered: Vaccine Safety Datalink Project*

Vaccine†	Frequency	Vaccine Combinations‡	Frequency
Oral polio (OPV)	732 652	Hep B alone	199 617
Diphtheria-tetanus-whole cell pertussis (DTP)	549 488	DTP + Hib + OPV	155 854
Haemophilus influenzae type b (Hib)	479 004	DTP + Hib + Hep B	104 099
Hepatitis B (Hep B)	455 746	DTP + OPV	99 501
Measles-mumps-rubella (MMR)	310 618	Hib + MMR	78 420
Combined DTP-Hib (DTPH)	147 650	DTPH + Hep B + OPV	60 100
Diphtheria-tetanus-acellular pertussis (DTaP)	126 982	DTaP + MMR + OPV	53 637
Influenza	27 014	DTP + MMR + OPV	49 333
Diphtheria-tetanus (DT)	16 944	DTP + Hib	45 256
Inactivated polio (IPV)	3 976	DTPH + OPV	44 670

* Analysis of first 3 years' data.
† Whether used alone or administered simultaneously with other vaccine(s).
‡ + between vaccine denotes simultaneous administration at different sites.

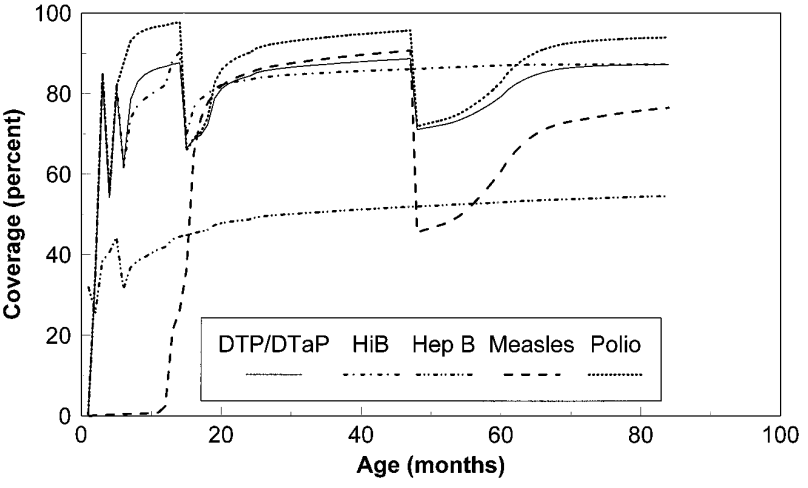

Fig 1. Composite vaccine coverage of Vaccine Safety Datalink Project cohort 1991–1993, by antigen and age, "up-to-date" status based on recommendations of the Advisory Committee on Immunization Practices.

when 80% of the cohort actually receives it increases for older children. The lowest coverage of routine childhood vaccinations in the cohort was for hepatitis B (universal recommendations began after the start of the study) and the second dose of measles vaccine (administered at primary school entry at two HMOs and secondary school entry at the others). The highest coverage was for polio vaccination, exceeding pertussis-containing vaccines by about 5% to 10%; this may measure the proportion of children with concerns about the safety of pertussis vaccines.

The simultaneous administration of multiple vaccines is very common. For example, when children younger than 6 months in our study received DTP vaccine, 97% of them also received *Haemophilus influenzae* type b (Hib) conjugate vaccine at the same visit. Thus, at certain ages it is difficult, if not impossible, to differentiate the possible effects of certain vaccine combinations. The vaccination schedules at the four HMOs are not identical, however, and this provides opportunities for separating the possible adverse effects of some vaccine combinations. For example, in the second year of life, one HMO gives DTP, Hib, oral polio vaccine, and measles-mumps-rubella (MMR) at the same visit, whereas the other three sites commonly give MMR and Hib at one visit and DTP (or diphtheria-tetanus-acellular pertussis [DtaP]) and oral polio vaccine at a later visit.

Analytical Studies

Other studies have shown that both DTP and MMR vaccines are associated with febrile seizures,[9,10,14] and this relationship was explored in the VSD study to illustrate its analytic capabilities. Because it is difficult to differentiate between febrile and other types of seizures without chart review, we analyzed all seizure types combined with the automated data. Using a Cox regression model stratified by HMO and birth date (children born in 3-day blocks), the relative rate of seizures in specific time windows after receipt of vaccine (same day, 1 to 3, 4 to 7, 8 to 14, and 15 to 30 days) was compared with periods before and more distant to the receipt of vaccine. Elevated rates of seizures, presumably mostly febrile, were found for both DTP and MMR, consistent with each vaccine's respective modes of biologic action (Fig 2). The risk varied for different periods after immunization. The risk of seizures after MMR, a live viral vaccine requiring a longer incubation period for viral replication, was delayed compared with that of DTP, a killed bacterial vaccine.

To examine the capability of the VSD cohort to differentiate the effects associated with simultaneous administration of vaccines, we examined the association between seizures and the administration of Hib and MMR vaccines. On crude analysis, a possible association was found 8 to 14 days after vaccination with both Hib (RR, 1.4; 95% confidence interval [CI],

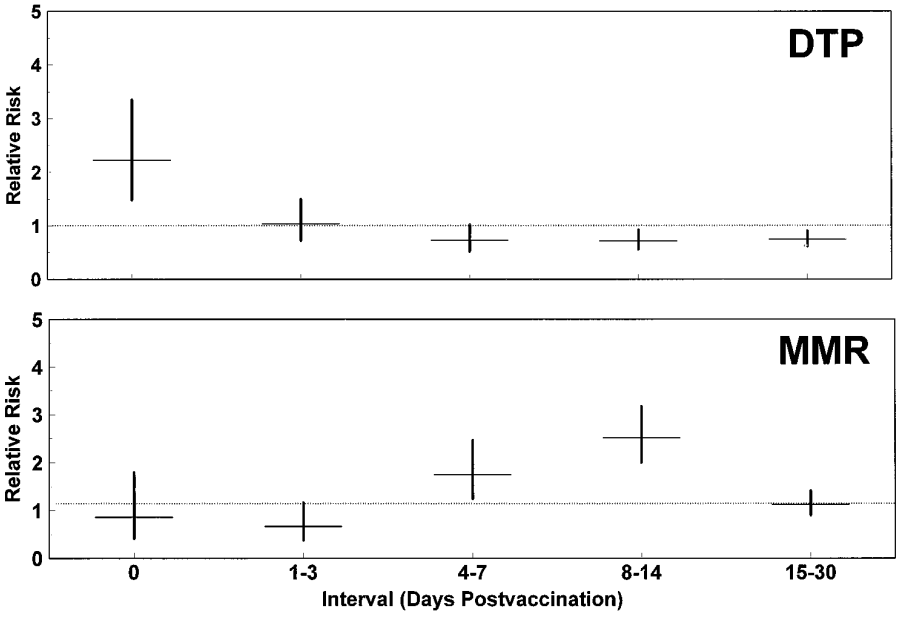

Fig 2. Relative risk and 95% confidence interval of seizures and persistent seizure disorders following DTP and MMR vaccinations in the Vaccine Safety Datalink Project. Risks are adjusted for simultaneous administration of other vaccines and based on automated screening data.

1.2 to 1.8) and MMR (RR, 2.3; 95% CI, 1.9 to 2.9) vaccines. Prior experience suggests that the association with Hib may be an artifact of the frequent coadministration of Hib with MMR during the second year of life and not a true association with Hib. Because the two vaccines are also frequently administered separately, we were able to adjust our analysis via regression to show that there is in fact no association between Hib and seizures (RR, 0.9; 95% CI, 0.7 to 1.2), whereas the association with MMR persisted (RR, 2.42; 95% CI, 1.8 to 3.2).

On screening of automated data, we found 8 to 14 days after immunization an apparent association between measles vaccine and invasive bacterial disease (RR, 2.3; 95% CI, 1.1 to 4.6). This association was not validated by chart review, because the exposed cases were found to have undergone evaluation of fever of unknown origin (possibly caused by the vaccine) but did not actually have culture-proven sepsis. This example shows that the ICD-9 codes selected for this outcome had the desired high sensitivity for surveillance purposes but needed validation.

Ad Hoc Studies

In addition to the planned vaccine safety studies, the infrastructure created by the VSD project permits timely investigation of new hypotheses. For example, when changes in vaccine policy are considered, it is often necessary to develop new information to evaluate such a potential change. The Advisory Committee on Immunization Practices recently considered lowering the age of the tetanus-diphtheria (Td) booster from 14 to 16 years of age to 10 to 12 years of age. However, information on the safety of the Td vaccine booster dose given to a younger age group was lacking. In the VSD study, we identified cohorts of 12 626 and 3379 children in the two age groups, respectively, who had received Td boosters. Comparison of rates of emergency department and hospital use within 7, 14, and 30 days after Td showed some differences between the two groups, which disappeared when visits for trauma and suture removal were excluded (Table 6), supporting the safety of this schedule change. This analysis further

TABLE 6. Rates of Hospitalization and Emergency Department Visits After Tetanus-Diphtheria (Td) Toxoid Vaccination by Age: Vaccine Safety Datalink Project

	10–12-y-Old		14–16-y-Old		Relative Risk	95% Confidence Interval	*P*
	(n)	(Rate)*	(n)	(Rate)*			
Td doses	3379		12 626				
Within 7 days of Td							
Hospitalization	9	139.0	22	90.9	1.53	0.65–3.48	.38
Emergency department visit	35	540.5	88	363.7	1.49	0.98–2.23	.06
Emergency department visit†	8	123.5	32	132.2	0.93	0.40–2.11	.99
Within 14 days of Td							
Hospitalization	13	100.4	33	68.2	1.47	0.74–2.90	.31
Emergency department visit	55	424.7	141	291.4	1.46	1.05–2.01	.03
Emergency department visit†	19	146.7	50	103.3	1.42	0.80–2.47	.25
Within 30 days of Td							
Hospitalization	18	64.8	44	42.4	1.53	0.85–2.72	.18
Emergency department visit	85	306.3	258	248.8	1.23	0.96–1.58	.11
Emergency department visit†	27	97.3	98	94.5	1.03	0.66–1.60	.96

* Rate per 1000 person-years.
† Excluding visits for trauma and/or suture removal.

illustrates the importance of adjusting for potential confounders in VSD studies. In this case, age itself affected rates of emergency department use independent of Td use.

DISCUSSION

Vaccines are generally administered to healthy persons, frequently infants and children. Therefore, the acceptable risk of adverse reactions to vaccines is lower than that for therapeutic agents for ill persons. This lower tolerable risk translates into the need to conduct studies to detect rare reactions (eg, attributable risks on the order of one per 10^5 to 10^6 doses).[4] Studies able to address such rare risks are possible only after licensure and general use, and they are large and expensive and may not provide conclusive results. For example, the National Childhood Encephalopathy Study was a case-control study that aimed to detect all hospitalizations in England and Wales for acute neurologic illness in children 2 to 35 months of age during a 3-year period.[5] The study findings were controversial, because the conclusion was based on only seven cases of chronic encephalopathy observed within 3 days of DTP vaccination.[32] These difficulties plus the limitations of passive surveillance[3,4] largely account for the relative sparsity of data on vaccine safety in the recent IOM review.[9,10]

In recent years, developments in health care organization, health information systems, and pharmacoepidemiologic methods have improved our capability to study rare drug reactions.[18] Walker et al[12] and Griffin et al[13] pioneered the use of such record linkage studies to evaluate vaccine safety. These studies were limited, however, by their relatively small sample sizes, retrospective design, and focus on the most severe reactions.[9] The VSD study attempts to overcome these shortcomings by prospective collection of vaccination, medical outcome, and covariate data under joint protocol at multiple sites. Selection of prepaid health plans also minimized potential biases resulting from data generated from fee-for-service claims. Substantial efforts have been required to implement accurate automated vaccination record systems for our cohort, which represents approximately 2% of the children in the United States. A list of key research questions and how best to answer them within the VSD has been elaborated. Quality control procedures and methodological approaches have also been developed. After all this development, does it work?

Although much remains to be done to improve the VSD, the early results are promising. Previously known associations between seizures and DTP and MMR vaccines have been validated in the prospective VSD cohort. This provides validation of the design, data collection, and analytic approaches of this project. The medical charts for many children with seizures identified from automated records are being abstracted. This will permit an evaluation of the accuracy of the automated system for a rigorous case-control analysis to distinguish between possible vaccine causation of first and subsequent seizures, as well as to characterize the types of seizures (eg, febrile versus other) associated with vaccinations.

In addition to studies to assess potential hypothesized vaccine associations, new ad hoc questions that arise from the VAERS, from changes in immunization schedules (eg, new vaccines such as varicella or the use of simultaneous vaccination) or from screening level cohort analyses in the VSD, can be addressed in a timely manner. For example, in response to concerns regarding a potential increased risk of arthropathy in adult women after rubella vaccination,[33] the VSD database was used to identify a cohort of women who had rubella immunity testing during pregnancy and their subsequent rubella immunization status. In a retrospective review of the women's charts, no association of any chronic arthritic condition was found with receipt of rubella vaccine.[34] The potential association between hepatitis B vaccination at birth and suspected neonatal sepsis work-up is being examined in another VSD study.

Each of the core databases created by the project (eg, vaccinations, medical outcomes, and potential covariates) has valuable applications. Each HMO in the VSD study uses its automated immunization records to set goals for improvement in vaccine coverage levels.[35,36] Children, clinics, and practices with inadequate immunization can be easily identified, and strategies can be developed to improve coverage. Research into barriers, missed opportunities, and recall systems for immunization have been performed.[37] Documentation for school-entry immunization requirements is also easily retrievable. These records have been used to calculate the level of childhood immunizations as of the second birthday included as quality measures in the Health Plan Employer Data and Information Set.[38,39] Finally, the VSD databases provide excellent bases for construction of broader regional immunization registries.[40]

Studies of many other pediatric illnesses via the VSD are also possible. Taking advantage of the cohort infrastructure created, plans are underway to expand the VSD study to examine: (1) vaccine safety issues in adolescents and adults, (2) the impact of vaccination programs on incidence of vaccine-preventable disease (eg, new varicella vaccine), (3) cost effectiveness of specific vaccines, and (4) safety and immunogenicity of new combined vaccine schedules in prospective phase II and III clinical trials nested within the cohort.[41]

The diversity in vaccination practice at the four HMOs and the clinic-to-clinic and day-to-day variations in practice permit useful contrasts in safety experiences. A study contrasting the safety of the second dose of measles vaccine administered at entry to primary school (as recommended by the Advisory Committee on Immunization Practices) versus entry to secondary school (as recommended by the American Academy of Pediatrics) is currently underway. As demonstrated by the Hib, MMR, and seizure example, the size of the VSD population may also permit separation of the risks associated with individual vaccines from those associated with vaccine combinations, whether given in the same syringe or simultaneously at different body sites. Such studies will be especially valuable in view of the new combined pediatric vaccines currently in development.[42]

Should the VSD study identify a vaccine reaction, data on attributable risk will be available, thereby permitting accurate risk-benefit assessment by both the public and policy makers.[43] Subgroup analyses may permit identification of risk factors, which may be useful in identifying contraindications to vaccinations. Research may then be launched to understand the pathogenesis of the reaction in these individuals, potentially leading to the development of safer vaccines. The incidence rates of reactions identified in the VSD should permit the evaluation and improvement of passive surveillance systems such as the VAERS. The VSD data will also be invaluable to other Public Health Service agencies sharing responsibility for vaccine safety. The Food and Drug Administration, in fulfillment of its regulatory responsibilities, is interested in potential product-specific differences in the vaccine safety profiles. The results of the VSD will also substantially enlarge the scientific basis for deciding whether to recommend compensation in alleged vaccine injury cases, fundamental to a fair and efficient vaccine injury compensation program.[44]

Amid these promises, a few caveats are appropriate. Although diverse, the population in the four HMOs currently in the VSD is not wholly representative of the US population in terms of geography or socioeconomic status. With current changes in health care organization and additional resources, it may be possible to broaden the scope of the VSD. In the interim, there is little reason to believe that these factors significantly influence the risk of vaccine reactions. More importantly, because of the high rate of vaccine coverage attained in the HMOs, few nonimmunized control subjects are available. The VSD must therefore rely predominantly on some type of risk interval analysis. The capability of this approach to assess associations between vaccination and adverse events with delayed or insidious onset (eg, autism) is limited. Similarly, the ability of the VSD to distinguish effects of combined or simultaneous vaccination fully may be limited should such practices become universal.

The VSD also cannot easily assess adverse events not currently captured in existing HMO databases, either because they do not result in health care consultations or because the data are not automated. Important nonautomated data sources relevant to the VSD study (eg, results of neurologic consultations) have required manual abstraction, coding, and computerization. The patient enrollment, health care practices, and health information systems at each HMO are dynamic, which may either aid or impede study of specific outcomes. Coding errors occurs inevitably in all data files to some extent, resulting in a decrease in our ability to detect a true association. The current VSD is also unable to examine the risk of extremely rare events after infrequent vaccinations, such as Guillain-Barré syndrome after each season's flu vaccine. Because the VSD relies on epidemiologic methods, it may not successfully control for confounding and bias in each analysis,[23] and inferences on causality may be limited. Finally, even if findings from the VSD may often be "negative" (ie, may show no elevations in risks in association with vaccination), one cannot absolutely "disprove" an alleged reaction.[10,45]

Despite these potential shortcomings, the VSD provides a new, essential, powerful, and cost-effective complement to our ongoing evaluations of vaccine safety in the United States. The capability of the VSD for reliable and consistent ascertainment should reassure the public of the adequacy of the surveillance for significant vaccine adverse events and the general safety of routine vaccine products. Enhanced public confidence is integral to maintaining or improving rates of vaccine acceptance at a time of rapid changes in vaccine schedules[46] and introduction of new vaccines.[16]

ACKNOWLEDGMENTS

We thank Drs Marie Griffin, Harry Guess, Edgar Marcuse, and Alec Walker for their invaluable insights as external advisors; the National Vaccine Program Office and the FDA for their financial assistance; and Bill Nichols, Phil Horne, and Dr Walt Orenstein for their unwavering support of the VSD project.

REFERENCES

1. World Bank. *World Development Report 1993: Investing in Health.* New York, NY: Oxford University Press; 1993
2. Expanded programme on immunization. Global advisory group—part 1. *Weekly Epidemiol Rec.* 1994;69:21–27
3. Chen RT, Rastogi SC, Mullen JR, et al. The Vaccine Adverse Event Reporting System (VAERS). *Vaccine.* 1994;12:542–550
4. Chen RT. Special methodological issues in pharmacoepidemiology studies of vaccine safety. In: Strom BL, ed. *Pharmacoepidemiology.* Sussex, United Kingdom: John Wiley & Sons; 1994
5. Miller DL, Alderslade R, Ross EM. Whooping cough and whooping cough vaccine: the risks and benefits debate. *Epidemiol Rev.* 1982;4:1–24
6. Krantz I, Taranger J, Trollfors B. Estimating incidence of whooping cough over time: a cross-sectional recall study of four Swedish birth cohorts. *Int J Epidemiol.* 1989;18:959–963
7. Kimura M, Kuno-Sakai H. Developments in pertussis immunisation in Japan. *Lancet.* 1990;336:30–32
8. Galazka AM, Robertson SE, Oblapenko GP. Resurgence of diphtheria. *Eur J Epidemiol.* 1995;11:95–105
9. Howson CP, Howe CJ, Fineberg HV, eds. *Adverse Effects of Pertussis and Rubella Vaccines: A Report of the Committee to Review the Adverse Consequences of Pertussis and Rubella Vaccines.* Washington, DC: National Academy Press; 1991
10. Stratton KR, Howe CJ, Johnston RB, eds. *Adverse Events Associated with Childhood Vaccines: Evidence Bearing on Causality.* Washington DC: National Academy Press; 1994
11. Schonberger LB, Bregman DJ, Sullivan-Bolyai JZ, et al. Guillain-Barré syndrome following vaccination in the national influenza immunization program, United States, 1976–1977. *Am J Epidemiol.* 1979;110:105–123
12. Walker AM, Jick H, Perera DR, et al. Diphtheria-tetanus-pertussis immunization and sudden infant death syndrome. *Am J Public Health.* 1987;77:945–951
13. Griffin MR, Ray WA, Fough RL, et al. Monitoring the safety of childhood immunization: methods of linking and augmenting computerized data bases for epidemiologic studies. *Am J Prev Med.* 1988;4(suppl):5–14
14. Walker AM, Jick H, Perera DR, et al. Neurologic events following diphtheria-tetanus-pertussis immunization. *Pediatrics.* 1988;81:345–349
15. Wassilak SGF, Glasser JW, Chen RT, Black S, Mullooly JP, Thompson RS. Design of a multicenter study of adverse events following vaccination in childhood. *J Clin Res Pharmacoepidemiol.* 1991;5:294. Abstract
16. Frontiers in medicine: vaccines. *Science.* 1994;265:1371–1404
17. Karch FE, Lasagna L. Adverse drug reactions. *JAMA.* 1975;234: 1236–1241
18. Strom BL, Carson JL. Use of automated databases for pharmacoepidemiology research. *Epidemiol Rev.* 1990;12:87–107
19. Farrington CP, Pugh S, Colville A, et al. A new method for active surveillance of adverse events from diphtheria/tetanus/pertussis and measles/mumps/rubella vaccines. *Lancet.* 1995;345:567–569
20. Roberts JD, Roos LL, Poffenroth LA, et al. Surveillance of vaccine-

related adverse events in the first year of life: a Manitoba cohort study. *J Clin Epidemiol.* 1996;49:51–58
21. Centers for Disease Control and Prevention. National Childhood Vaccine Injury Act. Requirements for permanent vaccination records and for reporting of selected events after vaccination. *MMWR.* 1988;37: 197–200
22. Griffin MR, Ray WA, Livengood JR, et al. Risk of sudden infant death syndrome following diphtheria-tetanus-pertussis immunization. *N Engl J Med.* 1988;319:618–623
23. Fine PEM, Chen RT. Confounding in studies of adverse reactions to vaccines. *Am J Epidemiol.* 1992;136:121–135
24. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am J Public Health.* 1992;92:703–710
25. Mullooly J. Misclassification model for person-time analysis of automated medical care database. *Am J Epidemiol.* 1996;144:782–792
26. Institute of Medicine. *Reliability of National Hospital Discharge Survey Data. Report of a Study.* Washington, DC: National Academy of Sciences; 1980
27. Fox LA, Stearns G, Imbiorski W, eds. *The Administrator's Guide to Evaluating Records and the Medical Records Department.* Chicago, IL: Care Communications; 1979
28. Mullooly JP, Black S, Thompson RS, Glasser JW, Chen RT, Rhodes P. Quality of large, linked HMO databases: estimation and correction of misclassification bias in screening analyses of childhood vaccine safety. *Pharmacoepidemiol Drug Safety.* 1994;3:S32. Abstract
29. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika.* 1986;73:1–11
30. Farrington CP, Nash J, Miller E. Case series analysis of adverse reactions to vaccines: a comparative evaluation. *Am J Epidemiol.* 1996;143: 1165–1173
31. Griffin MR, Ray WA, Mortimer EA, et al. Risk of seizures and encephalopathy after immunization with the diphtheria-tetanus-pertussis vaccine. *JAMA.* 1990;263:1641–1645
32. Marcuse EK, Wentz KR. The NCES reconsidered: summary of a 1989 workshop. *Vaccine.* 1990;8:531–535
33. Mitchell LA, Tingle AJ, Shukin R, Sangeorzan JA, McCune J, Braun DK. Chronic rubella vaccine-associated arthropathy. *Arch Intern Med.* 1993; 153:2268–2274
34. Ray P, Black S, Shinefield H, et al. Retrospective cohort evaluation of chronic rubella vaccine arthropathy. *Abstracts of the 34th Interscience Conference on Antimicrobial Agents and Chemotherapy.* Washington, DC: American Society for Microbiology; 1994:240. Abstract
35. Davis R, Vadheim C, Black S, et al. Immunization tracking systems: experience of the CDC vaccine safety datalink sites. *HMO Practice.* 1997;11:13–17
36. Payne T, Kanvik S, Seward R, et al. Development and validation of an immunization tracking system in a large health maintenance organization. *Am J Prev Med.* 1993;9:96–100
37. Lieu T, Black S, Sorel M, et al. Missed opportunities for immunization in a HMO. *Am J Public Health.* 1994;84:1621–1625
38. Lieu TA, Black SB, Sorel ME, Ray P, Shinefield HR. Would better adherence to guidelines improve childhood immunization rates? *Pediatrics.* 1996;98:1062–1068
39. Thompson RS, Taplin SH, McAfee TA, Mandelson MT, Smith AE. Primary and secondary preventive services in clinical practice. *JAMA.* 1995;273:1130–1135
40. Cordero JF, Guerra FA, Saarlas KN, eds. Developing immunization registries: experience from the All Kids Count Program. Am J Prev Med. 1997;13(suppl 1):1–128
41. Centers for Disease Control and Prevention. *Request for Proposals No. 200-95-0947(P). Comprehensive Linked Data Collection of Medical Events and Immunization (Vaccine Safety and Development Datalink).* Atlanta, GA: Centers for Disease Control and Prevention; 1995
42. Williams JC, Goldenthal KL, Burns DL, Lewis BP Jr, eds. *Combined Vaccines and Simultaneous Administration: Current Issues and Perspective.* New York, NY: New York Academy of Science; 1995
43. Hinman AR, Orenstein WA. Public health considerations. In: Plotkin SA, Mortimer EA, eds. *Vaccines.* Philadelphia, PA: WB Saunders; 1994: 903–932
44. Evans G. Vaccine liability and safety: a progress report. *Pediatr Infect Dis J.* 1996;15:477–478
45. Rothman KJ, ed. *Causal Inference.* Chestnut Hill, MA: Epidemiology Resources, Inc; 1988
46. Centers for Disease Control and Prevention. Recommended childhood immunization schedule—United States, 1997. *MMWR.* 1997;46:35–40